

ABSTRACT OF THE DISCLOSURE

1
2 To enhance the distribution or backup of data similar to previously distributed or backed up data,
3 to provide similarity testing, and other benefits, a first body of data T of size n is compressed
4 with respect to a second body of data S of size m *in-place*; that is, the memory containing S is
5 overwritten from left to right so that at no time is more than a total of $\text{MAX}\{m, n\} + O(1)$
6 memory used.